

IDS Bulletin

Transforming Development Knowledge

Volume 53 | Number 1 | February 2022

THEORY-BASED EVALUATION OF INCLUSIVE BUSINESS PROGRAMMES

Issue Editors **Giel Ton** and **Sietze Vellema**



Notes on Contributors	iii
Introduction: Contribution, Causality, Context, and Contingency when Evaluating Inclusive Business Programmes Giel Ton and Sietze Vellema	1
Systems, Sapiens, and Systemic Change in Markets: The Adopt-Adapt-Expand-Respond Framework Ben Taylor and Jake Lomax	21
Using Theory-Based Evaluation to Evaluate Systemic Change in a Market Systems Programme in Nepal Edward Hedley and Gordon Freer	43
Assessing the Contribution to Market System Change of the Private Enterprise Programme Ethiopia Giel Ton, Ben Taylor and Andrew Koleros	63
The Search for Real-Time Impact Monitoring for Private Sector Support Programmes Fédes van Rijn, Haki Pamuk, Just Dengerink and Giel Ton	87
Monitoring Systemic Change in Inclusive Agribusiness Sietze Vellema, Greetje Schouten and Marijn Faling	103
Assessing Contributions Collaboratively: Using Process Tracing to Capture Crowding In Marijn Faling	123
Understanding Behaviour Change in Theory-Based Evaluation of Market Systems Development Programmes Jodie Thorpe	141
Glossary	165

Introduction: Contribution, Causality, Context, and Contingency when Evaluating Inclusive Business Programmes*

Giel Ton¹ and Sietze Vellema²

Abstract The private sector has become an important partner in development interventions that aim to make market systems more favourable for smallholders and low-income consumers of food. How to evaluate these inclusive business programmes is the central theme of this *IDS Bulletin*. It presents real-world experiences of practitioners and academics using theory-based evaluation. This introductory article highlights the approaches and methods used to assess systemic change and provide learning for adaptive management. It acknowledges the limits to attributing outcomes to programmes alone and proposes a way to generalise about effectiveness where outcomes are highly contingent on a specific contextual embedding. The article explores the synergy of the iterative reflections on the theory of change, the analytical approach of realist evaluation, and the conceptualisation of changes in firms' practices as emerging from behaviour systems where the motivations, opportunities, and capabilities of firms are not equally distributed.

Keywords value chains, impact evaluation, market systems, realist evaluation, contribution analysis, theory-based evaluation.

1 Introduction

Theory-based evaluations (TBEs) are widely used to evaluate complex development programmes in dynamic environments. They reflect on the logic of the theory of change of programmes, and on the way that inputs are used for activities that translate into outputs, outcomes, and impact. Process evaluation focuses especially on the links between the inputs, activities and outputs. Impact evaluations help these reflections by verifying and qualifying the contribution claims at outcome and impact level. However, there are methodological limits to doing so.

© 2022 The Authors. *IDS Bulletin* © Institute of Development Studies | DOI: 10.19088/1968-2022.102



This is an Open Access article distributed under the terms of the Creative Commons Attribution Non Commercial 4.0 International licence (CC BY-NC), which permits use, distribution and reproduction in any medium, provided the original authors and source are credited, any modifications or adaptations are indicated, and the work is not used for commercial purposes.

The *IDS Bulletin* is published by Institute of Development Studies, Library Road, Brighton BN1 9RE, UK
This article is part of *IDS Bulletin* Vol. 53 No. 1 February 2022 'Theory-Based Evaluation of Inclusive Business Programmes'.

In this *IDS Bulletin*, we discuss experiences of practitioners and academics in finding doable and creative ways to conduct impact evaluations of inclusive business programmes in the domain of food and agriculture. Classic impact evaluation designs that start at baseline and proceed with a follow-up design to measure quantitative net effects in response to an intervention can be useful in a broader mix of methods in order to reflect on the theories of change (Vaessen, Lemire and Befani 2020) – but only where direct attribution of outcomes to the programme activities is possible and plausible (Ton, Vellema and Ge 2014). Attributing 'net effects' to a support programme is inappropriate for outcomes that lie beyond the sphere of influence. Other more reflexive theory-based approaches of impact evaluation are needed to assess effectiveness.

This introductory article discusses several experiences of impact evaluations that tried to develop these alternative approaches. It presents what was learnt from these evaluations about the tools and methods used to produce credible and actionable insights. It outlines an approach that combines the iterative reflection on the theory of change; the analytical approach of realist evaluation to explain the contextual embedding of change processes; and the conceptualisation of business strategies as behaviour that emerges from a complex system of incentives, and where motivations, opportunities, and capabilities are unequally distributed: some firms will benefit more than others. The article first reflects, in Section 2, on the literature that discusses inclusive business as an avenue for realising development outcomes in the domain of food and agriculture. Next, it focuses on the issues of contribution and causality (Section 3), and context and contingency (Section 4). The article ends with a reflection in Section 5 on how theory-based evaluation can become more meaningful for learning.

2 Inclusive business and development

Working with businesses to achieve development impacts has become an important strategy that is consistently featured in development policies and interventions proposed by international donor agencies, development organisations, and national governments (Heinrich-Fernandes 2016). Many of these programmes focus on the agricultural sector (Osorio-Cortes and Albu 2021). The *Rural Development Report 2021* of the International Fund for Agricultural Development (IFAD 2021) envisions a transformation of food systems driven by vast networks of mid-stream agri-food entrepreneurs, connected to small-scale farmers and consumers. The 2021 UN Food System Summit also confirmed the strong involvement of the private sector in achieving the Sustainable Development Goals.

Policies and intervention strategies labelled as value chain, market system, or food system approaches involve business partners in working towards a more inclusive development, where

smallholders are favourably integrated and low-income consumers have access to affordable and healthy food (Pouw, Bush and Mangnus 2019). We use the term 'inclusive business' programmes when we refer to these development approaches. The ideas about what inclusive business entails vary, and the reality of inclusive business is characterised by contingency and specificity (German *et al.* 2020). In general terms, inclusive business refers to including low-income communities in business chains (Likoko and Kini 2017), and it assumes that vulnerable, small-scale actors benefit through their integration into the (agri)business value chain (German *et al.* 2020; Schoneveld 2020).

There is a growing body of impact evaluations of inclusive business development interventions. The 2021 BEAM Evidence Review synthesises these in indicators aiming to compare their effectiveness and cost efficiency (Osorio-Cortes and Albu 2021). BEAM Exchange is a community of practice around the theme Building Effective and Accessible Markets. Also, the International Social and Environmental Accreditation and Labelling Alliance (ISEAL) has a repository of literature in the online Evidensia database³ that looks at the effectiveness of market-based coordination mechanisms, especially in tropical commodity chains.

At a more general level, the Donor Committee for Enterprise Development (DCED) tries to get harmonised indicators for private sector development programmes that help to reflect on the contribution of businesses to development impacts (DCED 2017). This continuing focus on comparison for accountability is, however, not always well aligned with the objective of learning from these interventions, due to the complex nature, multiple components, and different types of support activities that are being compared. A focus on the contextualised understanding of the unfolding change processes from which outcomes emerge may reveal why some groups are benefiting more than others from the support.

There is widespread support in the literature for increasing the coordination between the public and private sectors to develop new business models that have positive developmental outcomes. Scholars use different lenses to analyse these inclusive businesses. The value chain lens looks at standard setting, coordination, sourcing arrangements, input delivery, and service provision at sector level (Stoian *et al.* 2012). The value chain literature helps to find models and modalities that are conducive to pro-poor and sustainable development along the entire value chain (Ros-Tonen *et al.* 2019), from the upstream, smallholder producers to the downstream, 'bottom-of-the-pyramid' consumers (Maestre, Poole and Henson 2017).

A recent review of this literature emphasises the contextual embeddedness of interventions and proposes a processual perspective for exploring how inclusive business 'refashions'

relations and partnering among and across market actors and public actors (Schouten and Vellema 2019). This dynamic processual understanding is also present in literature on entrepreneurship. Entrepreneurs need to find modalities to connect and coordinate behaviour of multiple actors, ranging from smallholder farmers (Kangogo, Dentoni and Bijman 2021) and youth (Barzola *et al.* 2019) to community-based enterprises (Dentoni *et al.* 2018), in order to achieve a resilient business model.

A related strand of literature calls for recognition and appreciation of the activities of economic actors that operate in the middle of agri-food chains or food provisioning (Liverpool-Tasie *et al.* 2020; Reardon 2015). The rules and practices of businesses operating in between smallholder producers and urban markets importantly shape the terms of access to markets for farmers and micro-entrepreneurs, and influence the balance between exploitation and rewarding their inclusion. This body of literature shows that value chains, businesses, and markets cannot be treated as homogenous entities, and using these concepts – as we do in this article – has the danger of simplification.

The literature also shows heterogeneous effects of inclusive business. The claims of businesses that they contribute to inclusive business and desirable development outcomes are not self-evident. Development impacts are conditional on evolving change processes that emerge in diverse actor constellations, which operate in dynamic market and natural environments. Consequently, programmes and evaluators cope with moving targets in uncertain and unpredictable market and business environments, creating conditions of constant flux. Meanwhile, they need to contribute to previously agreed impact domains, such as food and nutrition security or environmental sustainability.

To be impactful, inclusive business programmes need multiple and overlapping interventions that are implemented in collaboration with multiple partners in a highly dynamic complex business environment. Therefore, to learn about their effectiveness, evaluations will need to find a way to acknowledge the co-existence of mutually constituting practices and unravel the interdependence and interaction among mutual causal processes.

Furthermore, private sector programmes and business practices are unlikely to be the sole contributing factors to observed development processes. The behaviours, practices, and rules associated with the intervention or business generate responses in a web of interdependence. Hence, the level of control over development impacts is limited, and having impact goes beyond the programmes' sphere of direct influence. A challenge for monitoring and evaluation of these programmes is how to account for the contingency and uncertainty that are inherent in such unfolding change processes in the broader system.

Moreover, evaluations of inclusive business programmes need to be attentive to signs of changes that affect and reshape rules and practices which underlie the nature of doing business: the issue of systemic change. Likely, it is only possible to assess systemic change in markets after some years, when the changed business models or value chain coordination modalities have matured and proven sustainable at scale.

This *IDS Bulletin* presents several experiences and approaches for monitoring and evaluating development outcomes and systemic change in inclusive business programmes. For example, Taylor and Lomax (this *IDS Bulletin*) propose to capture these systemic outcomes with the Adopt-Adapt-Expand-Respond (AAER) framework, originally developed in the Springfield Centre (Nippard, Hitchins and Elliott 2014). Monitoring the reactions of stakeholders and their behaviour in reaction to pilot interventions, the AAER framework captures unintended outcomes; the Adopt and Expand quadrants are more geared to the intended outcomes of an intervention. Vellema, Schouten and Faling (this *IDS Bulletin*) describe how they developed a tool to collect comparable data from more than 60 partnerships to capture early signs of systemic change. Based on a selection of theoretical frameworks, they typify three categories of outcomes to support implementers and partnership facilitators in noticing and valuing the effects of unfolding systemic change towards more inclusivity: the refashioning of the terms of inclusion of smallholders, the access of low-income consumers to food, and innovative leadership of the private sector that is doing 'business as unusual' (Vellema *et al.*, this *IDS Bulletin*). Hedley and Freer (this *IDS Bulletin*) use the Qualitative Impact Protocol (QulP) to scope for early signs of transformative change in markets.

The verification and critical assessment of contribution claims related to (early signs of) systemic change depend on the data that become available before, during, and after the implementation of the programme, the quality of the analysis and synthesis process, and the 'sense-making' about that evidence. Even when evaluation approaches can differ substantially in their ontological and epistemological assumptions (Stern *et al.* 2012), they all need data to generate the inputs for the sense-making process and causal inference process. The contributions by both practitioners and academics in this *IDS Bulletin* (see Table 1) respond to the call by Barbrook-Johnson *et al.* (2021: 5) to enlarge the toolbox available for complexity-appropriate evaluation.

We highlighted the methodological innovations in each of these articles that have potential to make impact evaluations more complexity-aware. At the same time, we acknowledge that the authors' accounts, including ours, of how the impact evaluations were designed and used are inevitably biased, as most of the authors operate 'in an environment where there are significant incentives to appear competent, minimise the

Table 1 Overview of the articles and programmes discussed in this introduction article

Authors	Name of programme	Nature of the interventions	Methodological innovations	
			Contribution and causality	Context and contingency
Taylor and Lomax	Generic	Market system development programmes	Nested and interlocked market systems	The Adopt-Adapt-Expand-Respond (AAER) framework captures ripple effects of pilot interventions in the wider market system
Hedley and Freer	Samarth-Nepal Market Development Programme (Samarth-NMDP)	Market system development in Nepal, especially in vegetable and dairy value chains	Contribution analysis through top-down and bottom-up research	Qualitative Impact Protocol (QulP) captures unbiased perceptions of impacts and change processes
Ton, Taylor and Koleros	Private Enterprise Programme Ethiopia (PEPE)	Market system development, especially in the leather, vegetable, and cotton sector, including labour sourcing in industrial parks	Interlinked research design, using firm-level surveys, process tracing case studies, and macro-economic modelling	Flexible results monitoring system using actor-based theories of change
van Rijn, Pamuk, Dengerink and Ton	Pioneering Real-time Impact Monitoring and Evaluation (PRIME)	Coaching and training of small and medium-sized enterprises to improve business management and export capacities	Online survey module to ask perceptions of impact and compare contribution scores on a range of outcomes	Real-time monitoring in a setting of dynamic navigation
Vellema, Schouten and Faling	2SCALE	Facilitating more than 60 partnerships for the (scaling of) inclusive agribusinesses fostering food and nutrition security in Africa	Using structured impact pathways embedded in each partnership for spelling out the sequential change processes	Contextualised monitoring of early signs of systemic change
Faling	Community Revenue Enhancement Through Agricultural Technology Extension (CREATE)	Linking smallholders to the barley and beer value chain in Ethiopia	Assessing pieces of evidence to verify a contribution claim, using process tracing	Supporting business partners in exploring their span of influence in a sector or industry
Thorpe	Developing Effective Private Education Nigeria (DEEPEN)	Improve the quality of education provided by private schools in Lagos	Using COM-B model to shift focus to behavioural change of firm	Graphical way to distinguish outcomes at different system levels using the COM-B model
	Financial Sector Deepening Trust Kenya (FSDK)	Generate sustainable livelihood improvements through better financial sector capacity and operations		
	Gender Transformative and Responsible Agribusiness Investments in South-East Asia (GRAISEA)	Improve livelihoods of women and men small-scale producers through more responsible and inclusive value chains and private sector investments		

problems, and to make things neater than the real messy process' (Rogers and Peersman 2014: 93). We remind the reader, therefore, to be cautious, and not to see any of the presented approaches and tools as the 'silver bullet' that resolves the challenges of the impact evaluation of inclusive business programmes.

Below, we elaborate major monitoring and evaluation (M&E) challenges and the possibilities to address these under two headings: **Contribution and causality** (Section 3), pointing to the co-existence of multiple intertwined causal processes; and **Context and contingency** (Section 4), pointing to the unpredictability and uncertainty of markets and the behaviour systems of those operating in these markets.

3 Contribution and causality

Lemire, Whynot and Montague (2019) give a nice overview of the increasing complexity of systems that are present in change processes. They present the spectrum of causal complexity in programme theories, where the simplest one is described as 'A leads to C', and the most complex one (the embedded-complex version) as 'A plus B leads to C because of D, under condition E'. Inclusive business programmes are clearly to be characterised as the latter, the embedded-complex ones; inclusive business programmes intentionally try to trigger changes that depend on other contextual conditions and incentives, apart from the support provided by them. A programme is (at most) a contributory factor in the process of generating inclusive business outcomes. Mackie (1974: 63) would call them an INUS factor – an 'insufficient but non-redundant part of a condition

Box 1 How to address INUS factors?

- **Insufficient factor:** Acknowledge that other conditions need to be in place for the programme support to work – a complex change process that the programme could not create alone.
 - **Non-redundant factor:** Verify whether the intervention is only 'accompanying' a causal process that would have been in place and created the outcome, without the support provided by the intervention playing any role of importance.
 - **Unnecessary configuration:** Acknowledge that other configurations (where the intervention does not take part) might exist that could also have resulted in the same outcome.
 - **Sufficient configuration:** Verify whether the outcome indeed happened at all and can be plausibly linked to the change process that has been supported.
-

which is itself unnecessary but sufficient for the result' (see Box 1). For an impact evaluation that follows a theory-based evaluation approach, we need contribution-verifying methods (did the intervention contribute to the change process?), and methods that reflect on the importance of this contribution (did it matter?). We consider these two questions in turn.

3.1 Did the intervention contribute to the change process?

Contribution-verifying methods need to show that the intervention is not redundant in a more complex change process that might not have taken place without the intervention. Firms develop new or refined service delivery models to include smallholders or to reach low-income consumers. That is the more complex change process, involving input providers, knowledge, and financial services, etc. The question in contribution-verifying research is whether the activities or resources of the support programme have played a non-redundant role in this change: would it have happened anyhow, without the support?

Faling (this *IDS Bulletin*) presents a nice example of process tracing as a way to critically verify whether a systemic change in the value chain of beer (crowding in of other malt factories) is indeed causally related to earlier support provided in the sector by an inclusive business programme. Though inspired by Bayesian reasoning (updating our confidence in the claim as more pieces of evidence become available), fortunately, Faling does not go as far as to compute the probability in a quasi-quantitative way, as recent process tracing literature suggests (Bennett, Charman and Fairfield 2021), and which Befani and Stedman-Bryce (2017) have coined as contribution tracing.

Bayesian updating starts with an estimate for the belief before a piece of evidence is considered (the prior probability), and results in an estimate that incorporates this new knowledge (the posterior probability). But, an informative evaluation of an inclusive business programme does not answer only one question – whether the programme is non-redundant in a change process – nor the probability that each arrow in the theory of change is true. Therefore, Bayesian updating might be a sophisticated but too narrow method to feed reflections about the effectiveness of inclusive business programmes.

However, what is clear is that a process tracing exercise implies a systematic process of seeking and critically assessing evidence. Faling (this *IDS Bulletin*) illustrates how process tracing can be used as a practical approach for explicating and scrutinising key assumptions in their contribution claims.

Hedley and Freer (this *IDS Bulletin*) piloted, among a wider set of top-down and bottom-up methods used, another tool to verify the contribution of the market development programme Samarth in Nepal. The tool, called Qualitative Impact Protocol (QuIP)

(Copestake, Morsink and Remnant 2019), implies that researchers talk with intended project beneficiaries about the main changes in their lives over a pre-defined recall period, and are prompted to share what they perceive as the main drivers of these changes, including to whom or what they attribute the change. The premise of QuIP is that the intended beneficiaries know a great deal about what has caused and affected changes in their lives, and what influenced their active decisions to start or stop doing certain activities (Copestake *et al.* 2019: 4). QuIP is explicitly not interested in inferring average effects but aims to explain or explore variation in the wellbeing outcomes.

Evaluators operate in an environment that creates strong incentives to look for confirming evidence (Rogers and Peersman 2014). QuIP is particularly keen to avoid this confirmation bias and is, therefore, often 'blindfolding' the researchers that do the interviews in a way that these do not know who commissioned the study or what support intervention is being evaluated. However, QuIP does not necessarily imply blindfolding. In the QuIP application in Samarth, documented by Hedley and Freer (this *IDS Bulletin*), the impact evaluation wanted to learn about the implementation modalities of service providers in value chains; they were less focused on assessing the outcomes of the changed value chain relation for the final beneficiaries (the smallholder producers). They learned in the process that evaluators needed to be sufficiently knowledgeable about the intricacies of the support and the heterogeneous effects that these modalities may have for different stakeholder groups in order to ask pertinent questions and unearth practical learning.

3.2 Did the contribution matter?

The market system development programmes discussed in this *IDS Bulletin* are all multi-year and well-resourced programmes with many activities, and are ready to respond to emerging issues and bottlenecks in the markets. Consequently, there will almost always be a clear (uncontested) non-redundant contribution to one or more intended outcomes, significant or insignificant as these outcomes might be. There will always be one or more firms that have changed their business models due to the increased market intelligence, capabilities, or opportunities that come with the support. Simply showing a contribution is, therefore, not enough. The question of the importance of this contribution needs to be addressed and is often the main reason why impact evaluations of inclusive business programmes are commissioned.

There are some principles and methods that help to reflect on the importance of contributions. Most importantly perhaps is that the commissioners of evaluations acknowledge that exact numbers are not needed but rough estimates suffice. For example, the Dutch Directorate-General for International Development⁴ (DGIS-RVO 2017) accept an estimate of the number of jobs or smallholders that were directly and indirectly supported by

an intervention, asking for estimates of reach instead of net effects. This opens the way for modelling the likely impacts of a programme according to several scenarios and based on explicit normative assumptions. A computable general equilibrium model, for example, could be used to estimate how the economy reacts to a plausible range of low and high growth rates in specific subsectors. The assumptions used in these extrapolations of effects are sensitive to normative decisions about the model parameters, especially when these models are built specifically for the evaluation.

Recently, the development finance institutions agreed upon a harmonised model, the Joint Impact Model (Steward Redqueen 2021). Even when inherently speculative and inexact, the use of a joint model would yield comparable indicators of development impact to judge the importance of programmes within and across countries. Instead, Mayne (2019a) argues that, rather than measure the size of the importance quantitatively, we need to ask the question: what is the relative importance of a specific causal factor in a wider configuration of factors? He suggests several ways of data collection that may help to (normatively) judge the importance of an intervention, after it is verified as being a non-redundant causal factor (Mayne 2019a: 5–6):

(1) the perceived influence of the causal factor in bringing about a change; (2) the role played by the causal factor in bringing about a change; (3) the funds expended by the causal factor; and (4) the magnitude of the constraint to change faced by the causal factor. In doing so, Mayne refrains from assessing the importance of a contribution in an objective, quantitative way.

Both van Rijn *et al.* (this *IDS Bulletin*) and Ton, Taylor and Kleros (this *IDS Bulletin*) assess the importance of a contribution in a subjective, quantitative way. They made use of a survey module that explicitly asked for the perception of the managers themselves about the importance of the contribution of the support. The survey tool was developed in response to Robert Chambers' recommendation made in the discussions about rigour in impact evaluation (Chambers 2009): 'If you want to know the impact – Ask them!'

Van Rijn *et al.* (this *IDS Bulletin*) applied the tool in an impact evaluation of two Dutch programmes that targeted managers of small and medium-sized enterprises in developing countries with coaching and capacity building. They asked the managers about their perception of change and the contribution of the support in eight areas of business management, when comparing themselves with similar firms in the sector.

Ton *et al.* (this *IDS Bulletin*) applied a similar tool to collect perception of change in perceived constraints in the institutional environment that affect business performance. The Private Enterprise Programme Ethiopia (PEPE) worked in several sectors

of the Ethiopian economy and aimed to trigger innovation processes that would improve sector performance. The online survey module does not directly mention PEPE but asks the managers about the influence of the relevant service providers or institutional arrangements that had received support from PEPE. The contribution score tool is based on a combination of two rankings of ordinal Likert scale answer categories, one in response to a question about the level of change in a reference period, and a second in response to a question about the importance of the specific support activity to these changes.

The combination of the two answer categories results in a ranking that reflects the importance of the contribution. The overall pattern of the contribution scores per area shows where the contribution is relatively important or less important. Using the INUS wording, the perception questions help to reflect on the non-redundancy and sufficiency of the intervention in change processes. Often, the change is there, but the respondent answers that the intervention played no role at all in the change process.

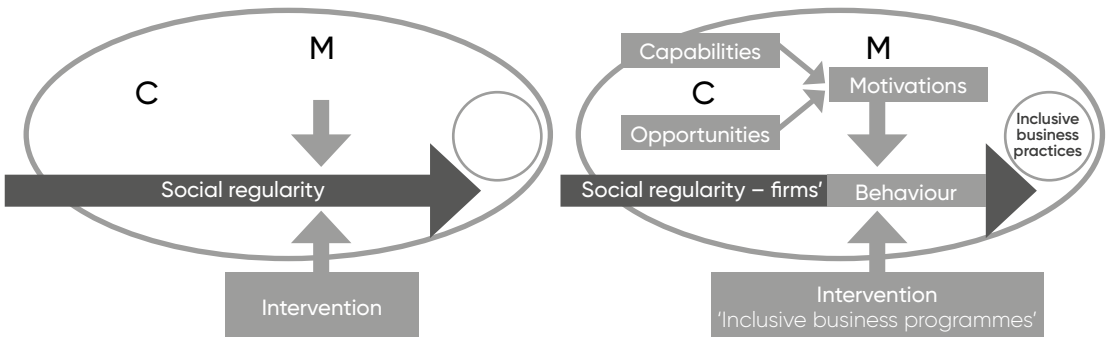
Of course, confirmation bias is still a threat to validity, but in both applications of the tool, a nuanced picture appeared with outcomes where more improvements were perceived, and with outcomes where the perception of change was less positive, which helped reflections about effectiveness of the portfolio of support activities on offer or in development.

These examples show that there are entrance points for improvement. The authors point to the need for more realism on the side of commissioners, requiring less precisely measured assessments of the ultimate outcomes and development impacts. Also, there needs to be (re)valorisation of the use of stakeholder perceptions, instead of observable outputs or outcomes to assess the importance of a programme's contributions to complex causal change processes, but acknowledging the bias that this could have. And finally, there is a need for the deployment of methods to critically assess alternative explanations, including the possibility that the intervention played no or only a marginal role in the change processes. With these 'ingredients' in the mix of methods, it proves possible to generate informative and credible impact evaluations of complex interventions like inclusive business programmes.

4 Context and contingency

The issue of contingency on context is highly relevant for inclusive business programmes, where market systems change due to the resources and reflection offered by the programme to the stakeholders involved – which subsequently triggers these stakeholders to do things differently. It is clear that the outcomes will always depend on the conditions in which these stakeholders are situated, and on the incentives and disincentives they are facing. The role of context is central to the realist evaluation approach,

Figure 1 Integrating COM-B in a realist CMO analysis of changing business practices



Note COM-B is 'capabilities, opportunities, and motivations for behaviour'; CMO is 'context, mechanisms, and outcome pattern'.

Source Authors' own, based on Pawson and Tilley (2006).

where the question 'What works, for whom, under what conditions, and why?' invites the evaluator to qualify in detail the exact role of an intervention in the change process and the conditions that are required for the intended change process to work.

In realist research, all causal mechanisms have a defined generalisation domain (Chen 1994), and need a reflection and analysis of the contextual conditions under which they are triggered (Pawson and Tilley 2006). The realist lens focuses on important components of programmes, acknowledges the heterogeneity of effects within these specific interventions, and identifies the configurations of conditions that enable these components to work well or explain why they fail.

A realist perspective translates into outlining multiple configurations of context, mechanisms, and outcomes. It is important to stress the configurational element, where 'it is not the ingredients that make the dish but how these are brought together in the cooking process' (Pawson and Manzano-Santaella 2012: 189). The object of the realist analysis will change depending on the type of support (resources and reasoning) and the scale where the activity is looking for outcomes. Each programme activity will have a specific outcome pattern, trigger particular mechanisms, and have specific contextual conditions.

However, at each scale level, the change processes can be conceptualised as changing a social regularity that results from a CMO-configuration – where CMO is context, mechanisms, and outcome pattern. The left-hand side of Figure 1 shows Pawson and Tilley's famous egg-shaped CMO-configuration (Pawson and Tilley 2006): an intervention changes the context C in a way that mechanisms M are triggered and change a social regularity resulting in an outcome pattern O.

Acknowledging that there is neither a silver bullet intervention nor a universal causal law in social systems, realists explore the generalisation domain of any conclusion about the 'what works' question. Accordingly, evaluation articulates and incorporates a selection of middle-range theories about the conditions under which an intervention might work elsewhere – and the mechanisms that would explain this. This is visible in how Vellema *et al.* (this *IDS Bulletin*) endeavour to unravel the composite nature of inclusive development and work towards categories of processual outcomes, reflecting the terms on which certain social groups are included in food provisioning. This is a step towards assessing whether the assumed achievement of 'business as unusual' materialises and, in realist terms, whether a social regularity is altered in a way that makes inclusive agribusiness rewarding for smallholder farmers and food affordable for low-income consumers. This helps to answer the 'what works, for whom' question.

To get closer to the actual mechanisms that generate change, a realist approach to impact evaluation of inclusive business programmes can opt to zoom in on choice-making and behavioural processes that generate changes in the form and substance of decision-making (Westthorp 2012, 2013) and subsequently refashion the practices and rules of individual firms or networks of firms. Following the conceptualisation of Michie, van Stralen and West (2011) and Mayne (2019b [2016]), firms or other organisational actors need to have the Capacities, the Opportunities, and the Motivations to make the Behaviour change (the COM-B model). Experiences with the new behaviour feed back into and change the capacities, opportunities, and motivations, as part of a structuration process where structures are reproduced and changed through agency (Giddens 1984). Inclusive business programmes have interventions and activities that aim to influence ongoing social processes (social regularity) by triggering the motivation of firms (mechanisms) to change their business models (behaviour) in a way that markets become better, fairer, and more inclusive (outcome pattern).

Especially for micro-level change in firms, where the outcome is a change in practices, the COM-B model of behaviour change might be useful to operationalise this realist analysis (see Figure 1). The decision-making of firms and farmers involves multiple, often competing, motivations and incentives that emerge due to certain conditions, and changes in these conditions. For example, most business-oriented development programmes discussed in this *IDS Bulletin* do not provide financial resources but knowledge, coaching, or capacity building.

These programmes address the capabilities and motivations of firms. However, in doing so, they experience that even when firms have the right motivation and capability, there must also be the opportunity to enact their behaviour. Thorpe (this *IDS Bulletin*)

presents a graphic representation using the COM-B model to illustrate the findings of three impact evaluations, with critical elements at the macro and meso levels that drive behaviour at the micro level. Ton *et al.* (this *IDS Bulletin*) apply it when they discuss experiences in the PEPE programme, where actor-based theory of change models (Koleros and Mayne 2019) are used to detail intervention strategies and guide the building of a results-based monitoring system (Yohannes 2020; Posthumus *et al.* 2020).

The combination of realist thinking and the conceptualisation of behaviour as a COM-B system offers a fruitful way to develop (nested) theories of change and middle-range theory. This may smooth the path for a learning-oriented approach to impact evaluation that recognises that achieving intended outcomes is contingent on how the interventions configure with conditions.

Some mechanisms (motivations) that lead to inclusive business outcomes are only triggered by inclusive business programmes under the right conditions. Likewise, some mechanisms (motivations) may be triggered that explain why interventions have no results. Accordingly, realist analysis generates actionable insights on the limits of interventions. The realist focus on causal mechanisms-in-context proves particularly relevant when the main role of a programme is piloting, innovating, and experimenting with intervention modalities that are expected to be scaled or implemented in other contexts.

Realist evaluators aim to develop actionable middle-range theories (Cartwright 2020; Pawson and Tilley 2006), around the question 'What works for whom, under what conditions, and why?' The message is that there are no silver bullets – there are no universal laws in social science – but all outcomes are context-dependent and often contingent on many complex, intertwined mechanisms, incentive structures, and motivations that mean that there is a high level of contingency, serendipity, and surprise involved.

5 Making evaluation meaningful

Most impact evaluations of inclusive business or market system programmes use a theory-based evaluation approach (Osorio-Cortes and Albu 2021). This raises the question, 'Whose theories are selected, combined, and refined?' A good theory-based evaluation asks for a critical engagement with a plural set of theories. Smart data collection and sharp analysis and synthesis alone are not enough. The evaluation process and outputs also need to be informative for the stakeholders involved.

Several authors in this issue (van Rijn *et al.*; Ton *et al.*; Hedley and Freer) conclude that more interaction and sense-making between implementers and evaluators is needed. Although we acknowledge that there are more stakeholders than implementing agencies involved in an impact evaluation, the

learning by the implementing agency was for most authors in this *IDS Bulletin* the main goal. But the authors also show that this is the stakeholder group for whom they struggled most to prove the usefulness and value of systematic and rigorous forms of monitoring and evaluation.

Under the right conditions, the presented approaches and tools might work and accelerate the learning loops for adaptive management. Three conditions appear as necessary components in the causal configurations that result in a high-quality theory-based evaluation: (1) interested 'listeners' as the audience of the evaluation, especially the commissioners and implementing agency; (2) rigour in anticipating and addressing validity threats to the conclusions derived from the methods used; and, last but not least, (3) sufficient resources for an appropriate mix of methods.

The inclusive business programmes in this *IDS Bulletin* provide examples of theory-based evaluation approaches that go beyond the tick-box exercises that still characterise large parts of the monitoring and evaluation field. The authors piloted processes and generated outputs that were meant to be functional for learning, and especially the comparative learning about effectiveness and relevance of intervention modalities across a portfolio of supported projects. Vellema *et al.* (this *IDS Bulletin*) discuss critical issues in the set-up of such a learning-oriented evaluation system within the 2SCALE programme that fits the navigation of partnership in dynamic contexts. The PRIME evaluation, described by van Rijn *et al.* (this *IDS Bulletin*), developed tools for real-time monitoring in an organisational setting where a monitoring, evaluation, and learning (MEL) system already existed that needed to be upgraded to meet external reporting requirements.

Other impact evaluations, such as PEPE (Ton *et al.* this *IDS Bulletin*; Koleros 2020; Johannes 2020) and Samarth (Hedley and Freer, this *IDS Bulletin*) are embedded in the evaluation system of the donor that uses annual reviews and external impact evaluations with baseline, midterm, and endline data collection and synthesis (ICAI 2015), and where the evaluators need to ensure that learning occurs 'within these predictive management standardised tools and templates' (Koleros 2020: 63). However, all experiences presented in this *IDS Bulletin* acknowledge that it is not easy to find ways to make learning useful for commissioners and implementing agencies. The actual use of findings depends on many factors that are beyond the control of evaluators.

Notes

- * We thank the contributors to the *IDS Bulletin* for their openness and reflexivity while sharing the lessons learned from designing and implementing evaluation tools in the dynamic setting of inclusive business programmes. This article benefited from the critical review by Dominic Glover and the contributing authors. Publication of the *IDS Bulletin* was supported by the 2SCALE programme.
- 1 Giel Ton, Research Fellow, Institute of Development Studies, University of Sussex, UK.
- 2 Sietze Vellema, Associate Professor, Knowledge, Technology and Innovation group, Wageningen University, the Netherlands.
- 3 For more information see **Evidensia**.
- 4 In full, Ministerie van Buitenlandse Zaken-Rijksdienst voor Ondernemend Nederland [Dutch Ministry of Foreign Affairs-Netherlands Enterprise Agency]

References

- Barbrook-Johnson, P.; Castellani, B.; Hills, D.; Penn, A. and Gilbert, N. (2021) 'Policy Evaluation for a Complex World: Practical Methods and Reflections from the UK Centre for the Evaluation of Complexity across the Nexus', *Evaluation* 27.1: 4–17
- Barzola, C.L. *et al.* (2019) 'Challenges of Youth Involvement in Sustainable Food Systems: Lessons Learned from the Case of Farmers' Value Network Embeddedness in Ugandan Multi-Stakeholder Platforms', in R. Valentini, J.L. Sievenpiper, M. Antonelli and K. Dembska (eds), *Achieving the Sustainable Development Goals Through Sustainable Food Systems*, Cham: Springer
- Befani, B. and Stedman-Bryce, G. (2017) 'Process Tracing and Bayesian Updating for Impact Evaluation', *Evaluation* 23.1: 42–60
- Bennett, A.; Charman, A.E. and Fairfield, T. (2021) '**Understanding Bayesianism: Fundamentals for Process Tracers**', *Political Analysis*, DOI: 10.1017/pan.2021.23 (advance online publication) (accessed 16 November 2021)
- Cartwright, N. (2020) 'Middle-Range Theory: Without It What Could Anyone Do?', *Theoria: An International Journal for Theory, History and Foundations of Science* 35.3: 269–323
- Chambers, R. (2009) 'Perspectives on Impact Evaluation: Approaches to Assessing Development Effectiveness', video, Cairo: International Initiative for Impact Evaluation (3ie)
- Chen, H-T. (1994) *Theory-Driven Evaluations*, Thousand Oaks CA: SAGE
- Copstake, J.; Morsink, M. and Remnant, F. (2019) *Attributing Development Impact: The Qualitative Impact Protocol Case Book*, Rugby: Practical Action
- DCED (2017) *The DCED Standard for Measuring Achievements in Private Sector Development: Control Points and Compliance Criteria*, London: Donor Committee for Enterprise Development
- Dentoni, D.; Pascucci, S.; Poldner, K. and Gartner, W.B. (2018) 'Learning "Who We Are" By Doing: Processes of

- Co-Constructing Prosocial Identities in Community-Based Enterprises', *Journal of Business Venturing* 33.5: 603–22
- DGIS-RVO (2017) *15 Methodological Notes – Instructions for Calculation, Validation and Reporting of Performance Indicators*, The Hague: Ministerie van Buitenlandse Zaken-Rijksdienst voor Ondernemend Nederland [Dutch Ministry of Foreign Affairs-Netherlands Enterprise Agency]
- German, L.A.; Bonanno, A.M.; Foster, L.C. and Cotula, L. (2020) '“Inclusive Business” in Agriculture: Evidence from the Evolution of Agricultural Value Chains', *World Development* 134: 105018
- Giddens, A. (1984) *The Constitution of Society: Outline of the Theory of Structuration*, Berkley CA: University of California Press
- Heinrich-Fernandes, M. (2016) *How to Create an Enabling Environment for Inclusive Business? Learning from Experience and Debates in Private Sector Development*, Business Environment Working Group Paper, London: Donor Committee for Enterprise Development
- ICAI (2015) *DFID's Approach to Delivering Impact*, Report 45, London: Independent Commission for Aid Impact
- IFAD (2021) *Transforming Food Systems for Rural Prosperity: Rural Development Report 2021*, Rome: International Fund for Agricultural Development
- Kangogo, D.; Dentoni, D. and Bijman, J. (2021) 'Adoption of Climate-Smart Agriculture Among Smallholder Farmers: Does Farmer Entrepreneurship Matter?', *Land Use Policy* 109: 105666
- Koleros, A. (2020) 'Accountable for Adaption: How Independent Evaluation Can Support Adaptive Programming within Bureaucracies', in B. Perrin and T. Tyrrell (eds), *Changing Bureaucracies: Adapting to Uncertainty, and How Evaluation Can Help*, New York NY: Routledge
- Koleros, A. and Mayne, J. (2019) 'Using Actor-Based Theories of Change to Conduct Robust Contribution Analysis in Complex Settings', *Canadian Journal of Program Evaluation* 33.3: 292–315
- Lemire, S.; Whynot, J. and Montague, S. (2019) 'How We Model Matters: A Manifesto for the Next Generation of Program Theorizing', *Canadian Journal of Program Evaluation* 33.3: 414–33
- Likoko, E. and Kini, J. (2017) 'Inclusive Business – A Business Approach to Development', *Current Opinion in Environmental Sustainability* 24: 84–88
- Liverpool-Tasie, L.S.O. et al. (2020) 'A Scoping Review of Market Links Between Value Chain Actors and Small-Scale Producers in Developing Regions', *Nature Sustainability* 3: 799–808
- Mackie, J.L. (1974) *The Cement of the Universe: A Study of Causation*, Oxford: Oxford University Press
- Maestre, M.; Poole, N. and Henson, S. (2017) 'Assessing Food Value Chain Pathways, Linkages and Impacts for Better Nutrition of Vulnerable Groups', *Food Policy* 68: 31–39

- Mayne, J. (2019a) **Assessing the Relative Importance of Causal Factors**, Centre for Development Impact (CDI) Practice Paper 21, Brighton: Institute of Development Studies (accessed 16 November 2021)
- Mayne, J. (2019b [2016]) 'The COM-B Theory of Change Model', Working Paper V5, unpublished
- Michie, S.; van Stralen, M.M. and West, R. (2011) 'The Behaviour Change Wheel: A New Method for Characterising and Designing Behaviour Change Interventions', *Implementation Science* 6: 42
- Nippard, D.; Hitchins, R. and Elliott, D. (2014) **Adopt-Adapt-Expand-Respond: A Framework for Managing and Measuring Systemic Change Processes**, Springfield Centre Briefing Paper, Durham: Springfield Centre (accessed 9 September 2021)
- Osorio-Cortes, L.E. and Albu, M. (2021) *The Results Achieved by Programmes that Use the Market Systems Development (MSD) Approach: A Narrative Synthesis of Current Evidence*, BEAM Evidence Review 2021, London: BEAM Exchange
- Pawson, R. and Manzano-Santaella, A. (2012) 'A Realist Diagnostic Workshop', *Evaluation* 18.2: 176–91
- Pawson, R. and Tilley, N. (2006) *Realist Evaluation, Realistic Explanation: Linking Explanatory and Interpretive Accounts in Applied Social Research*, Utrecht and Wageningen: CERES
- Posthumus, H.; Shah, R.; Miehlandt, A. and Kessler, A. (2020) *A Pragmatic Approach to Assessing System Change*, Durham: Springfield Centre
- Pouw, N.; Bush, S.R. and Mangnus, E.P.M. (2019) 'Editorial Overview: Inclusive Business for Sustainability', *Current Opinion in Environmental Sustainability* 41: A1–A4
- Reardon, T. (2015) 'The Hidden Middle: The Quiet Revolution in the Midstream of Agrifood Value Chains in Developing Countries', *Oxford Review of Economic Policy* 31.1: 45–63
- Rogers, P.J. and Peersman, G. (2014) '**Developing a Research Agenda for Impact Evaluation in Development**', *IDS Bulletin* 45.6: 85–99, DOI: [10.1111/1759-5436.12115](https://doi.org/10.1111/1759-5436.12115) (accessed 16 November 2021)
- Ros-Tonen, M.A.F. et al. (2019) 'Conceptualizing Inclusiveness of Smallholder Value Chain Integration', *Current Opinion in Environmental Sustainability* 41: 10–17
- Schoneveld, G.C. (2020) 'Sustainable Business Models for Inclusive Growth: Towards a Conceptual Foundation of Inclusive Business', *Journal of Cleaner Production* 277: 124062
- Schouten, G. and Vellema, S. (2019) 'Partnering for Inclusive Business in Food Provisioning', *Current Opinion in Environmental Sustainability* 41: 38–42
- Stern, E. et al. (2012) *Broadening the Range of Designs and Methods for Impact Evaluation*, London: Department for International Development
- Steward Redqueen (2021) **Joint Impact Model – Version 1.6** (accessed 11 November 2021)

- Stoian, D.; Donovan, J.; Fisk, J. and Muldoon, M. (2012) 'Value Chain Development for Rural Poverty Reduction: A Reality Check and a Warning', *Enterprise Development and Microfinance* 23.1: 54–69
- Ton, G.; Vellema, S. and Ge, L. (2014) '**The Triviality of Measuring Ultimate Outcomes: Acknowledging the Span of Direct Influence**', *IDS Bulletin* 45.6: 37–48
- Vaessen, J.; Lemire, S. and Befani, B. (2020) *Evaluation of International Development Interventions: An Overview of Approaches and Methods*, Washington DC: World Bank
- Westhorp, G. (2013) 'Developing Complexity-Consistent Theory in a Realist Investigation', *Evaluation* 19.4: 364–82
- Westhorp, G. (2012) 'Using Complexity-Consistent Theory for Evaluating Complex Systems', *Evaluation* 18.4: 405–20
- Yohannes, L. (2020) *Adaptive Management: From the Inside Looking Out*, Addis Ababa: Enterprise Partners

This page is intentionally left blank